LOCAL AND GLOBAL PREDICTION ALGORITHMS: THE EFFECTIVENESS IN SHORT/ LONG RANGE FORECASTING OF EXTREME METEOROLOGICAL EVENTS IN THE INDIAN SUBCONTINENT

SHORT SYNOPSIS

Submitted in partial fulfillment of the requirement of the degree of

DOCTOR OF PHILOSOPHY

by

Rashi Aggarwal
Registration no: MRUCSDN012K210053

Under the supervision of

Prof. (Dr.) Manpreet Kaur
Professor
Department of Computer Science & Technology, Manav Rachna University, Faridabad, Haryana

Dr. K.C Tripathi
Associate Professor
Maharaja Agrasen Institute of Technology, Rohini, Delhi

Department of Computer Science & Technology
School of Engineering

MANAV RACHNA UNIVERSITY
Sector-43, Aravali Hills, Faridabad

Year 2025

MANAV RACHNA UNIVERSITY

Declared as State Private University Vide Haryana UGC Act 26 of 2014

MANAV RACHNA UNIVERSITY

Sector-43, Aravali Hills, Faridabad

Date: 16.5.25

DECLARATION BY THE SCHOLAR

I hereby declare that the short synopsis entitled **"LOCAL AND GLOBAL PREDICTION ALGORITHMS: THE EFFECTIVENESS IN SHORT/ LONG RANGE FORECASTING OF EXTREME METEOROLOGICAL EVENTS IN THE INDIAN SUBCONTINENT"** being submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy in the Department of Computer Science & Technology under the School of Engineering of Manav Rachna University, Faridabad during the academic year 2024-25.

It is the work that is expected to be carried out under the supervision of Prof. (Dr.) Manpreet Kaur and co-supervision of Dr. K.C Tripathi.

Further, it is declared that the short synopsis is within the plagiarism limit and copy of the same is attached as per the University norms.

Rashi Aggarwal
Registration Number: MRUCSDN012K210053
ORCID ID: https://orcid.org/0000-0003-2358-3512

MANAV RACHNA UNIVERSITY

MANAV RACHNA UNIVERSITY
Sector-43, Aravali Hills, Faridabad

Date: 16.5.25

CERTIFICATE BY SUPERVISOR(S)

This is to certify that, the short synopsis entitled **"LOCAL AND GLOBAL PREDICTION ALGORITHMS: THE EFFECTIVENESS IN SHORT/ LONG RANGE FORECASTING OF EXTREME METEOROLOGICAL EVENTS IN THE INDIAN SUBCONTINENT"** is an original piece of work done by research scholar Ms. Rashi Aggarwal under the supervision of Prof. (Dr.) Manpreet Kaur and co-supervision of Dr. K. C Tripathi for the degree of Doctor of Philosophy in the Department of Computer Science & Technology under the School of Engineering of Manav Rachna University, Faridabad during the academic year 2024-25.

Further, it is certified that the short synopsis is within the plagiarism limit and copy of the same is attached as per the University norms.

Manpreet Kaur

Signature of Supervisor
Prof. (Dr.) Manpreet Kaur
Professor
Department of Computer Science & Technology
School of Engineering
Manav Rachna University
Faridabad, Haryana

Signature of Co-Supervisor
Dr. K.C Tripathi
Associate Professor
Department of IT
MAIT
GGSIP University
Rohini, Delhi

ABSTRACT

Owing to its impact on human life worldwide, weather forecasting has drawn the interest of several experts from diverse scientific fields. Due to the dynamic nature of the atmosphere, the prediction of atmospheric parameters suffers from two major limitations: (i) sources of error arising out of wrong initial conditions or poor understanding of the problem and (ii) their inability to model the peaks. There are two types of mathematical models to address these limitations- statistical and dynamical. The dynamical meteorologists have attempted to minimize the sources of error and got mixed results. This leaves a research gap so far as statistical models are concerned. The problem of modelling the peaks or in other words identifying the regions of error spikes by general formulation is not settled by either the dynamical or the statistical approach. Thus (i) and (ii) above form the basis of the present study. This study is aimed at investing the statistical models such as regression, neural network and Super-ensemble Neural Network for modelling the All India or Regional Summer monsoon rainfall using time series signals or external forcing of Antarctic sea ice concentration (AnSIC) or Indian Ocean Sea Surface Temperature (SST). The study investigates two questions pertaining to statistical meteorological models- (i) whether or not the methods of error minimization in dynamical models can be applied successfully in statistical models and (ii) if the statistical models are as effective at modelling the peaks as they are in predicting the general points of the time series.

Keywords: Neural networks, ensemble neural network, Super-ensemble neural network, AnSIC

Table of Contents

1. Introduction 6
   1.1 Introduction to weather prediction 6
   1.2 Motivation for the present study 12
   1.3 Problem formulation 14
2 Literature Review 16
3 Description of Broad Area 24
4 Objectives of the Study 25
5 Methodology to be adopted 26
   5.1 The system 26
   5.2 The predictors and forcing 26
   5.3 The models 26
   5.4 The analyses of results 27
6 Expected Outcome of the Research 27
7 References 27

1. Introduction

1.1 Introduction to weather prediction

Any equation serves as a streamlined abstraction of a real-world process, often constructed with the help of computer simulations. In the field of meteorology—the discipline that examines atmospheric and oceanic behavior—forecasting techniques have undergone major transformation. Historically, approximately a hundred years ago, weather forecasting primarily depended on analyzing prior climate data. This approach, referred to as extrapolation, aimed to project upcoming weather by identifying trends from previous patterns.

Today's most sophisticated forecasting tools are Global Circulation Models (GCMs) [1]. The technique of Numerical Weather Prediction (NWP) through Global Circulation Models (GCMs) is commonly referred to as dynamical modeling. This method relies on advanced computational algorithms to solve differential equations that describe the fundamental physical principles of the atmosphere and climate system.

1.1.1 Dynamical modelling (Numerical Weather Prediction, NWP)

Various mathematical aspects of Global Circulation Models (GCMs) have been explored and refined over time [2]. Dynamical modeling, which is another term for Numerical Weather Prediction (NWP) models, involves solving actual equations to predict future values of weather parameters. These equations are typically differential equations that describe the physical laws governing the atmosphere and oceans.

Back in the early 1920s to mid-1920s, Hoskins and Hodges were pioneers in this field. They used calculators to solve these equations and make weather predictions, which was a significant step forward at the time [3]. However, it wasn't until the 1950s, with the advent of computers and advancements in observational technology, that this field really took off.

It became clear that even for very local and short-term forecasts, large-scale mathematical computations were essential. This realization drove the development of more sophisticated models and computational methods, leading to the highly accurate

and detailed weather predictions we have today. Formally, a NWP model can be represented mathematically as

$$\Delta\vec{W} = M\,(\vec{W}^{initial})\Delta t \tag{1}$$

where $\Delta\vec{W}$ represents change in weather, Δt represents the time duration for the change and M is a function that takes $\vec{W}^{initial}$, a set of weather parameters as input, to calculate the change, $\Delta\vec{W}$, in weather. The process represented by (1) is known as model integration. Actual implementation of this model is a complex task and requires meteorological expertise along with huge computational resources. M is the mathematical (NWP) model that solves the governing differential equations. If the model M is changed, the differential weather, $\Delta\vec{W}$, changes. Finally,

$$\vec{W}^{forecast} = \vec{W}^{initial} + \Delta\vec{W} \tag{2}$$

The principle around which the NWP models work is that by providing (i) correct initial conditions i.e. correct values of $\vec{W}^{initial}$ and (ii) correct governing equations i.e. correct model M, accurate prediction of the future weather states are possible using (1). This is illustrated in fig. 1.

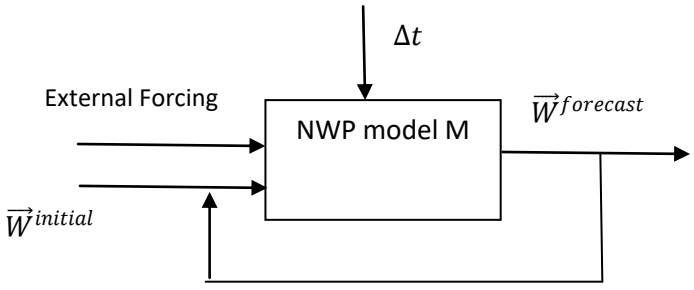


Figure 1: Working of a typical NWP model

None of the conditions, (i) or (ii) listed above are met even today and can never be accurately met in future owing the following errors:

Type 1 error related to incorrect initial conditions:

i. Correct initial conditions means the least count error of measuring instruments is zero which is practically impossible,
ii. Correct initial conditions means the floating point and round off errors in digital computations are zero which is again impossible

Type 2 error related to incorrect governing equations:

i. 100% knowledge of the governing equations which is again impossible owing to the local and global factors.

Any of the above errors, type 1 or type 2 will generate incorrect value of $\vec{W}^{forecast}$. This error quickly multiplies (see fig. 1) at future steps and after a few iterations the predictability falls.

However, researchers have tried to minimize these errors. Type 1 error is minimized by taking several initial conditions each resulting from a perturbation applied to actual readings. If $\vec{P}$ represents the set of actual parameters to be supplied as input to a model and $\vec{P}_i$ represents the i^{th} measurement of $\vec{P}$ then the actual input supplied to the model is

$$\hat{P} = \frac{1}{N}\sum_i \vec{P}_i \tag{3}$$

where N is the total number of measurements taken for $\vec{P}$. This reduces the errors due to inaccurate observations. Another method is to take the average of model outputs obtained from simulating the system N times with input $\vec{P}_i$. If $\vec{S}_i$ represents the output with $\vec{P}_i$ as input, the final output is

$$\widehat{M} = \frac{1}{N}\sum_i \vec{S}_i \tag{4}$$

Equations (3) and (4) represent a modelling technique called the ensemble forecasting.

Type 2 error, due to limited knowledge of the full system, was attempted to be resolved by [4][5] using multi-model ensemble (MME) forecasting which essentially is described as:

$$\hat{F} = V_j \widehat{M}_j \tag{5}$$

where $\hat{F}$ is the final output, $\widehat{M}_j$ is the j^{th} model the output of which is described by (4) and V_j is the weight associated with output of each model. This technique of using

different models for prediction has been successful as it has been demonstrated that MME forecasts are better than the ensemble forecasts [6].

1.1.2 Statistical modelling (Statistical Weather Prediction, SWP)

Statistical modelling refers to identifying patterns in the data to infer about general characteristics which can be used for prediction [8]. Basis of statistical methods for atmospheric sciences has been found in various studies [7][9]. The objective behind statistical and dynamical model is to make predictions for the advancement of socio-economic conditions of public at large. Whereas the dynamical modelling is essential for the understanding of the problem the statistical modelling is required for making quick models that can be easily deployed for making predictions. A statistical model such as the regression does not require knowledge of governing equations as (1) above. Such a model essentially identifies patterns in the data. This is generalized in the following equations

$$\hat{y} = f(\vec{x}) \tag{6}$$

where $\vec{x}$ is the independent variable on which the parameter y depends, $\hat{y}$ is the estimated value of y. f is the functional relationship between $\vec{x}$ and y obtained by a statistical model usually by minimizing a cost function of the form

$$J = \sqrt{\sum_i (y_i - \hat{y_i})^2} \tag{7}$$

where the subscript i refers to data point number. Equation (7) forms the basis of statistical modelling same way as equation (1) is the basis of numerical modelling. The distinction made is that while in (1) M is a specific NWP model, J in equation (7) is a general cost function to computer f in equation (6). However, the model M can be considered a general model for sequence reconstruction if it is customized to the specific sequence in question. The f in (6), however, is only dependent on the data i.e., $\vec{x}$. If the data refers to meteorology, then f is an approximation to the meteorological relationship between $\vec{x}$ and y. In meteorology, the predictor $\vec{x}$ must 'lag' by some steps behind the predictand y. This is called number of steps of prediction. Thus, customizing (6) for meteorology,

$$\hat{y}_n = f(\vec{x}_{n-k}) \tag{8}$$

where k is the lag value implying that prediction is made k steps ahead. What is common in (1) and (8) is that the data consumed is same. The $\vec{f}$ in (6) is M in (1) if both are perfect. However, a major difference is that unlike in NWP model (fig. 1) there is no feedback connection in a SWP model. In other words $\vec{x}_{n-k}$ is a vector of observational data and has no component of model output at a previous stage. Thus, an error introduced at any step will, not affect the forecast at subsequent steps. Some similarities and differences in (1) and (8), i.e., methodologies of NWP and SWP are summarized in table 1.

The major difference is the absence of feedback loop in SWP. This is the major deviation in the forecasting methodology. The advantages and disadvantages of one type of forecasting model over the other are listed in table 2. The two key arguments are (i) both NWP and SWP models give similar forecasts and (ii) the statistical models are not suited for developing an understanding of the dynamics of the system but sometimes, can be used for critically examining the present states and future behaviors. It is this aspect that shall be used in the present study to investigate a very vital area of meteorological research i.e. error dynamics.

Table 1: NWP and SWP, a comparison of methodologies

S. No	NWP (Eq. 1 & 2)	SWP (Eq. 6 & 7)
1	M is a *numerical weather model* that solves numerically, the equations that calculate $\Delta\vec{W}$ from $\vec{W}^{initial}$.	f is a general regression model which is fed with $\vec{x}$ to statistically estimate y (denoted by $\hat{y}$) using the cost function J. For weather forecasting $\vec{x}$ is same as $\vec{W}^{initial}$ and $\hat{y}$ same as $\vec{W}^{forecast}$.
2	The progression from $\vec{W}^{initial}$ to $\vec{W}^{forecast}$ (eq. 1) is incremental as illustrated in fig. 1. An error	In equation (8) the value of $\vec{x}_{n-k}$ is observed value and not derived from earlier step of prediction. Thus, an error introduced at a

	introduced at n^{th} step affects the prediction at subsequent steps.	step will not affect the forecast value at a subsequent step.
3	NWP model integrates the system numerically using differential equations and hence requires huge computational resources.	Prediction is made simply by making use of functional relationship obtained by minimizing a cost function which depends only on data values and not the intrinsic relation among them.

Table 2: NWP and SWP, a comparison of merits and demerits

S. No	**Parameter**	**NWP**	**SWP**
1	Computing resources	Requires huge computing resources to solve equation (1) for global and local equations numerically.	Requires fewer computing recourses because it is only equation (7) that needs to be solved. The determination of f is then straightforward.
2	Computation time	Requires more computation time to solve the equations numerically.	Requires comparatively less computing time as minimization of cost function in equation (7) is not iterative and is done based on the data.
3	Knowledge of the system	Requires full understanding of the dynamical laws, equation (1) and (2) governing the system.	Requires no understanding of the dynamical laws to solve equation (7). Only statistical formulation is needed.
4	Application scope	Are required for better understanding of the meteorological relations	Do not help in understanding the dynamics of the system. However, these may provide an insight into the future

		and future prediction of events	behavior of the system by careful analysis of the results.
5	Quality of forecast	Significantly better than a decade ago	Similar to that of NWP.
6	Predictability	An infinitesimal error may grow out of bounds after few iterations and predictability may fall.	Absence of feedback poses no or little such threat.

1.1.3 Dynamical aspects of statistical predictions

Although, as highlighted above (table 2, para 6) the statistical models make predictions based on lead lag correlation (eq. 8). The input at present state is independent of the output of previous stage and hence the errors introduced at any stage does not, in principle, multiply in subsequent stages. This leads to an important conclusion regarding the growth and decay of errors in various temporal regions of a dynamical system [9]. It was concluded (though for a toy model, Loremz) that predictions made by statistical models (Artificial Neural Networks and Farmer-Sidorowich model) are independent of regions of dynamical error growth and decay and as such statistical models are better suited for making predictions where the errors tend to grow. Another important characteristic of a dynamical system is the generation of peaks. Any prediction model, NWP or SWP, is appreciated if it can model the peaks rather than making average forecasts. Peaks in a time series may or may not be a property of the system. It may be a result of dynamics of the system or random local forces. Thus, modelling the peaks become difficult by any prediction model. This situation raises a question for research community:

Can a model generate peaks at the same locations where the system dynamics tend to generate them?

1.2 Motivation for the present study

We put up the question raised in the previous section here again- Can a model generate peaks at the same locations where the system dynamics tend to generate them?. We go on to elaborate on this question in this section and in the next section.

We have seen in the introductory section that for any natural process there are two broad techniques for modelling and hence making predictions- the NWP and the SWP. The SWP models are typically the regression models. We have also compared the two and enlisted the limitations of each. They complement each other and one doesn't pose a counter threat to the other.

In meteorological forecasting, there are areas of concern which are related to the system and not to the method of forecasting. We enlist two such concerns below i.e. (i) the problem of modelling the peaks and (ii) sources of errors.

1.2.1 The problem of modelling the peaks

It was put up in section 1(c) if a model can generate peaks at the same locations where the system dynamics tend to generate them. Here, a model can be statistical or dynamical. Our study is limited to statistical models only but in future the problem of modelling the peaks may be taken up by dynamical meteorologists to provide greater insight into the problem.

1.2.2 The sources of errors

Refer to type 1 and type 2 errors discussed in section 1(a) (also mentioned in table 2 (6)). Such errors are characteristics of NWP models. The SWP is not an iterative procedure (except for ANN models) and hence an error obtained at any step does not affect the error at any subsequent step. Since the initial condition problem is not related to SWP and also since SWP models don't require and understanding of the dynamical system, therefore, the procedures developed to reduce forecast errors in NWP (eq. 5) owing to these are not needed for SWP. However, for SWP models like the ANN [10] the model initialisation makes sense. ANNs are statistical models which use iterative learning algorithms for doing classification and regression. These are highly dependent on initial weights. Hence different initializations lead to different forecasts unlike the

regression models which make the same forecast irrespective of the number of times we use the model. This aspect of ANNs compare with the NWP. Thus, the moderations developed in equation (5) apply in the context of ANNs. But how well these moderations apply in the context of ANN? This is investigated in the present study.

We now proceed to formulate the problem from the above two aspects namely (i) the problem of modelling the peaks by SWP and (ii) the effect of initialisation and design in a class of statistical models such as the ANNs. For mathematical formulation we shall consider an arbitrary sequence and the problem of mathematically determining the peaks. Then we proceed to examine the theory for meteorological systems.

1.3 Problem formulation

A vital problem in any simulation/ modelling research, wherein a sequence is required to be reconstructed accurately is that of modelling the peaks and/ or the points of inflection. Surely the sequence is generated by empirical rules and is a non-deterministic one else the need for modelling is rendered useless.

1.3.1 The spatial and temporal regions of error spikes:

Time series of a meteorological parameter can be considered a sequence generated by empirical rules. This rule may be the association of the parameter with an external forcing. The extrema in the sequence may be considered as extreme weather event. Thus, the predictability of the extrema in the sequence is equivalent to the predictability of the extreme weather event associated with that parameter.

The prediction errors give vital information on the predictability of a time series of a parameter. A model may perform well during normal events but fail to model the extreme weather events. This makes the model less useful as it is the extreme event that is more desirable to be predicted.

For the purpose of quantification of the problem, the following mathematical notations are used:

S : The time series/ sequence of an atmospheric/ Oceanic parameter

N : Number of points in S

R : The geographical region under consideration.

M : The set of maxima in the sequence$S$.

P : The set of minima in the sequence$S$.

$\vec{f}$: The set of external forcing that generated S.

F : Statistical model for the prediction of S, $M$and P using external forcing as $\vec{f}$.

$\hat{S}$: The model generated sequence

$\widehat{M}$: The model generated set of maxima

$\hat{P}$: The model generated set of minima

The objective of model F is to make $\hat{S}$ as close to S as possible, i.e. .

$$\hat{S} \to S \tag{9}$$

However, for the model to make acceptable prediction in the extreme weather region (spatial or temporal), we must also have:

$$\widehat{M} \to M \tag{10}$$

$$\& \quad \hat{P} \to P \tag{11}$$

Remember that (9) does not imply (10) and (11).

Describe the cost function needed to calibrate F while modeling S :

$$J = \left[\frac{1}{N}\sum_{i=1}^{N}(s_i - \hat{s}_i)^2\right]^{1/2} \tag{12}$$

The cost incurred on the prediction of M and P may be similarly defined as J_M and J_P respectively. The vital question is if

$$J \sim J_M \ \& \ J \sim J_P \tag{13}$$

An affirmative answer requires that prediction errors while predicting normal events in a meteorological time series are same as those while predicting extreme weather events.

1.3.2 Sources of error and their elimination:

So far as initial condition errors are concerned, dynamical models tend to reduce these by taking initial conditions perturbed from the actual input and then taking the average of all outputs. This is called ensemble forecasting. To take care of (ii), Multi-Model

Ensemble (MME) **[8]** forecasting technique was introduced. This can be summed up as:

$$M = \frac{1}{(j*i)} \sum_i M_i \sum_j M_i^j \quad (14)$$

where M_i^j is the output of the j^{th} initialisation of model M_i. M_i represents the average output of j initialisations of model M_i. M is the average output of all such outputs with all models.

Whereas the dynamical modelling has attained success in MME techniques there is no parallel explanation in statistical (ANN) modelling to account for these errors. We develop the class of ANN models, The super-ensemble models, in line with the MME models and investigate the prediction accuracy.

2 Literature Review

During the late Indian Summer Monsoon, extreme rainfall events have been connected to changes in Arctic Sea ice. Researchers have also investigated the impact of Antarctic Sea Ice on All over and regional rainfall of Indian Subcontinent [14]. Various dynamic and statistical tools, resulted in good prediction analysis [15][17][20].

The JJAS months patterns can be forecasted over long periods, primarily using past rainfall data, without the need for considering distant climate events. [23].

To better understand the complex relationship between Indian Summer Monsoon Rainfall (ISMR) and Sea Surface Temperature (SST), the researchers used Artificial Neural Network (ANN) models. Average (Ensembles) are favored as they capture the non-linearity in the data and reduce both bias and variance in predictions [21][24][25].

Based on point and interval forecasts, various homogeneous neural network averaging methods were used. It was shown that RMSE was further decreased when seasonality component was removed from data series [19].

The Elman recurrent neural network, radial basis function network, multilayered perceptron network, Hopfield model for hourly temperature, wind speed, etc., are among the artificial neural network ensembles used to forecast the weather. Ensemble approaches were used to build predictions based on forecasting. Ultimately, the temporal complexity decreased since

the nonlinear feed forward neural network ensemble showed more promise in terms of accuracy and capacity. [15].

Using a temporal convolutional network (TCN) to forecast the Indian Ocean's monthly SSTs, long-term SST can be reliably predicted offering a foundation for the investigation of significant oceanic physical processes. This approach effectively addressed the challenge of making precise long-term and large-scale sea surface temperature (SST) predictions. Furthermore, its performance in extracting spatial and temporal properties surpassed that of the CNN and CFSv2 models [11].

Several methodologies for constructing super ensembles have been explored in the literature. In this study, the ensemble is formed by aggregating a number of ANNs, each initialized with distinct parameters. This approach leverages the diversity of individual models to enhance predictive accuracy and robustness [18]. This finding highlights the model's potential for improving climate forecasts and understanding the complex interactions between polar regions and global weather patterns. Table further demonstrates the literature:

Year	Author	Title	Journal	Major findings
2024	Akash Verma, Leena, Sachin Budakoti	Understanding the dynamics of 2024 extreme heat event in India: Spatial variability, hydrometeorological impacts, and model evaluation	Atmospheric Research	The study showed how vegetation and the 2024 heatwave are related. Due to model constraints and the complexity of the tropical system, the Noah Model's forecasting ability drastically declined after two days.
2023	Ammar Mohammed, Rania Kora	A comprehensive review on ensemble	‘Journal of King Saud University –	The aggregated learning can aid

		deep learning: Opportunities and challenges	Computer and Information’	improve generalization across unseen data.
2022	Tianying Sun , Yuan Feng 1, Chen Li and Xingzhi Zhang	High Precision Sea Surface Temperature Prediction of Long Period and Large Area in the Indian Ocean Based on the Temporal Convolutional Network and Internet of Things	Sensors	It was suggested to use a temporal convolutional network (TCN) to forecast the Indian Ocean's monthly SSTs.The research proposed that the TCN model can reliably forecast the long-term SST, providing a foundation for the study of significant oceanic physical processes.[11]
2022	Yiming V.Wang, Thomas Larsen, Stefan Lauterbach, Nils Andersen, Thomas Blanz, Uta Krebs-Kanzow, Paul Gierz,and Ralph R. Schneider,	Higher sea surface temperature in the Indian Ocean during the Last Interglacial weakened the South Asian monsoon	PNAS	This study demonstrated that, in addition to solar insolation, internal climate feedbacks are crucial for the variability of the South Asian monsoon during warm climate regimes. This work

				can help improve future climate model projections by highlighting the importance of understanding monsoonal rainfall controls under interglacial boundary conditions.[12]
2019	Joseph Nzau Mutemi	Improving Seasonal Climate Forecasts over Various Regions of Africa Using the Multimodel Super ensemble Approach	Atmospheric and Climate Sciences	The supersensemble approach for multimodels was introduced. The method made greater use of the expertise in the climate forecast data sets from atmosphere-ocean general circulation models that are operated at numerous global centres, such as the WMO Global Producing Centres (GPCs). [13]

2021	Chatterjee, S., Ravichandran, M., Murukesh, N. et al.	A possible relation between Arctic sea ice and late season Indian Summer Monsoon Rainfall extremes		It suggests that while natural climate variabilities in mid-latitudes may have played a role in historical trends, recent changes in the Arctic climate are more strongly influenced by human-induced warming, potentially affecting the relationship between SIE and ISMR. [14]
2019	Pandey, A.C., Das, I.M.L., Rai, S., Mishra, A.P., Pandey, V.K., Dwivedi, S., Mitra, A. & Tripathi, K. C.	Mathematical Modeling of Atmosphere and Ocean Processes Around Antarctica, in Atmospheric Research in Antarctica: Present Status and Thrust Areas in Climate Change	Taylor and Francis	This paper implemented Semtner and Winton's three-layer sea ice models to simulate sea ice thickness and snow surface temperatures for polar regions. Temperature, geo-potential height,

				and horizontal wind vectors have been predicted around Indian Antarctic Station Maitri using polar MM5. SST anomalies in SIO and ACC regions have also been simulated using artificial neural networks (ANNs).[15]
2018	Mukherjee, S., Aadhar, S., Stone, D. & Mishra, V	Increase in extreme precipitation events under anthropogenic warming in India	Weather and Climate extremes	This paper presented computational models from the CMIP5 and C20CÍ detection and assignation tasks to evaluate changes in the observed climate, the effect of anthropogenic warming on severe rainfall events, and expected changes in the frequency of excessive

				precipitation in India. [16]
2018	K.C. Goudaa, S. Nahaka, P. Goswamib	Evaluation of a GCM in seasonal forecasting of extreme rainfall events over continental India	Weather and Climate Extremes, Elsevier	Anomalies in Antarctic sea ice areas have been simulated using artificial neural networks (ANNs). Several dominating cycles are found using FFT, and the impact of both their presence and absence on trend modelling is evaluated. [17]
2017	Shruthy Govindan, K. C. Tripathi	Superensembling of Artificial Neural Network Models for Investigating the Effect of Polar Sea Ice on Indian summer Monsoon Rainfall	International Journal of Computer Applications	The Superensemble model, which is based on multiple ANN models' ensembles, is defined in this study. An array of many ANNs with different initializations makes up the ensemble's ANN model. Super ensembles have been approached in

				a variety of ways. [18]
2011	Iffat A. Gheyas, Leslie S. Smith	A novel neural ensemble architecture for time series forecasting.	NeuroComputing, ScienceDirect	This research proposes a new symmetric neural network ensemble technique called Generalised Regression. Neural network ensemble: a synthesis of various machine learning techniques for single-variate time series forecasting. The algorithms were evaluated using RMSE and point and interval predictions. [19]
2009	IML Das, KC Tripathi	Tele-Connection between Antarctic sea ice concentration and All India Rainfall	Proceedings of International Symposium on snow and Avalanches	In this paper, Linear Regression & ANN models are used to predict AIR anomalies from AnSIC [20]
2008	P. Guhathakurta and M. Rajeevan	Trends in the rainfall pattern over India	International Journal of Climatology,	In contrast to simulation crop modelling, this

			Wiley InterScience	article showed that machine learning (ML) may produce reasonable forecasts more quickly and with greater flexibility. A committee of models (machine learning ensembles) can, nevertheless, outperform a single machine learning model since they can better represent the underlying distribution of the data and minimise prediction bias, variance, or both. [21]

3 Description of Broad Area

In meteorology, a particular interest is the modelling of the Indian summer monsoon rainfall [22] which, along with the variations in the Indian Ocean Sea Surface Temperatures (SST) [11][12][23] is the motivation behind most of meteorological studies in the Indian Ocean region. There are many extreme weather events that are the outcomes of these two phenomena.

These extreme events are nothing but extreme points in the time series of the parameter under consideration.

The study has three dimensions- (i) meteorological (ii) mathematical and (iii) statistical and machine learning. The aspect is the application of statistical prediction models to meteorological events (Indian monsoon rainfall) from the perspective of modelling the peaks. The SWP and the NWP models are discussed in the introductory section. The mathematical conceptualization of the problem of rare events is discussed in section 3.

The mathematical formulation is necessary to visualize the problem from general mathematical perspective and also to leave a scope to apply the solution to other disciplines as well. For example the results obtained herein can be used in other areas of application such as civil infrastructure where modelling the cracks and strength of material is a sensitive task given the stake involved.

4 Objectives of the Study

a. Identification of extreme local events of interest, such as floods, draughts and Indian Ocean dipole formation etc. Correlation analyses for the identification of suitable external forcing, (like the SST, polar sea ice concentration etc) that may have empirically generated the sequence.
b. Development of statistical models, F- like regression, moving averages, ANN and Super-ensemble ANN models that may be used to exploit the correlations in (a) for predicting the sequence S in general and extreme events M/P in particular.
c. Investigating the impact of external forcing $\vec{f}$ (in particular the polar sea ice concentration/ area) in the generation of S.
d. Carrying out statistical analyses of the model outputs to
    i. Explain the predictability of anomalous points in the sequence S leading to the predictability of rare weather events in the Indian subcontinent using statistical prediction models such as regression and Artificial Neural Networks (ANN).

ii. Investigate the models for removing the sources of error when using the Super-ensemble ANN model

5 Methodology to be adopted

As outlined in section 3 the study is aimed at (i) analysing the predictability of extreme events and (ii) identifying and removing the sources of error in forecasting the Indian and regional summer monsoon rainfall using the AnSIC and other predictors with ML based models.

5.1 The system

As outlined in the introduction section the present study aims to explore the system of meteorological parameter (s) particularly the Indian summer monsoon rainfall or the regional Indian rainfall for predictability in extreme weather events. The forcing shall be provided by the Sea Surface Temperature (SST) or the Antarctic Sea ice concentration (AnSIC). The basic question addressed is whether the extreme weather event, modelled as an anomalous point in the time series of the parameter is inherently predictable. For this we shall be using the local and global prediction algorithms.

5.2 The predictors and forcing

i. The time series of the parameter at appropriate lag
ii. The AnSIC/ SST forcing at appropriate lag and at appropriate spatial domain.

5.3 The models

i. The linear and non linear regression models (single and multiple)
ii. The Artificial Neural Network Models (ANN)
iii. The local prediction models such as the moving average
iv. The neural network model

5.4 The analyses of results

The results shall be analysed with the following objectives:

v. The predictability of the general points in the time series of the parameter
vi. The predictability of the extreme weather points in the times of the parameter modelled as anomalous point in the time series. The anomaly shall be determined by the degree of deviation from the mean or on the basis of analytical results
vii. The categorical forecast skill scores

6 Expected Outcome of the Research

The study is aimed at identifying whether an extreme event is a consequence of the dynamics of the system or a random process. It will be determined if an outlier point comes from the relationship between the variables or comes as a chance point. This will be calculated by calculating the errors in the normal time and the outlier time. if the error in the outlier time increases then it will mean that this point was a chance point. we will do this for all such points. finally, we may conclude the reason for the outlier point. this will help in making future models

7 References

[1] Lynch, P. (2008). The origins of computer weather prediction and climate modeling. *Journal of Computational Physics, 227*(7), 3431–3444. https://doi.org/10.1016/j.jcp.2007.02.034

[2] Richardson, L. F. (1922/2007). *Weather prediction by numerical process*. Cambridge University Press. https://doi.org/10.1017/CBO9780511618291

[3] Hoskins, B. J., & Hodges, K. I. (2002). New perspectives on the Northern Hemisphere winter storm tracks. *Journal of the Atmospheric Sciences, 59*, 1041–1061.

[4] Palmer, T. N., & Coauthors. (2004). Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bulletin of the American Meteorological Society, 85*, 853–872. https://doi.org/10.1175/BAMS-85-6-853

[5] Krishnamurthy, V., & Goswami, B. N. (2000). Indian monsoon–ENSO relationship on interdecadal timescale. *Journal of Climate, 13*, 579–595. https://doi.org/10.1175/1520-0442(2000)0132.0.CO;2

[6] Tripathi, K. C., Rai, S., Pandey, A. C., & Das, I. M. L. (2008). Southern Indian Ocean SST indices as early predictors of Indian summer monsoon. *Indian Journal of Marine Sciences*, *37*(1), 70–76.

[7] Krishnamurti, T. N., & Coauthors. (2001). Real-time multianalysis-multimodel superensemble forecasts of precipitation using TRMM and SSM/I products. *Monthly Weather Review, 129*, 2861–2883.

[8] Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (3rd ed.). Academic Press.

[9] Mittal, A. K., Singh, U. P., Tiwari, A., Dwivedi, S., Joshi, M. K., & Tripathi, K. C. (2015). Short-term predictions by statistical methods in regions of varying dynamical error growth in a chaotic system. *Meteorology and Atmospheric Physics*. https://doi.org/10.1007/s00703-015-0375-3

[10] Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford University Press.

[11] Sun, T., Feng, Y., Li, C., & Zhang, X. (2022). High precision sea surface temperature prediction of long period and large area in the Indian Ocean based on the temporal convolutional network and Internet of Things. *Sensors, 22*(4), 1636. https://doi.org/10.3390/s22041636

[12] Wang, Y. V., & Larsen, T. (2022). Higher sea surface temperature in the Indian Ocean during the Last Interglacial weakened the South Asian monsoon. *Proceedings of the National Academy of Sciences*, *119*(10). https://doi.org/10.1073/pnas.2107720119

[13] Mutemi, J. N. (2019). Improving seasonal climate forecasts over various regions of Africa using the multimodel superensemble approach. *Atmospheric and Climate Sciences, 9*(4). https://doi.org/10.4236/acs.2019.94038

[14] Chatterjee, S., Ravichandran, M., Murukesh, N., et al. (2021). A possible relation between Arctic sea ice and late season Indian summer monsoon rainfall extremes. *npj Climate and Atmospheric Science, 4*, 36. https://doi.org/10.1038/s41612-021-00191-w

[15] Pandey, A. C., Das, I. M. L., Rai, S., Mishra, A. P., Pandey, V. K., Dwivedi, S., Mitra, A., & Tripathi, K. C. (2019). Mathematical modeling of atmosphere and ocean processes around

Antarctica. In *Atmospheric Research in Antarctica: Present Status and Thrust Areas in Climate Change*. Taylor & Francis.

[16] Mukherjee, S., Aadhar, S., Stone, D., & Mishra, V. (2018). Increase in extreme precipitation events under anthropogenic warming in India. *Weather and Climate Extremes, 20*, 45–53. https://doi.org/10.1016/j.wace.2018.03.005

[17] Gouda, K. C., Nahak, S., & Goswami, P. (2018). Evaluation of a GCM in seasonal forecasting of extreme rainfall events over continental India. *Weather and Climate Extremes, 21*, 10–16. https://doi.org/10.1016/j.wace.2018.05.001

[18] Govindan, S., & Tripathi, K. C. (2017). Superensembling of artificial neural network models for investigating the effect of polar sea ice on Indian summer monsoon rainfall. *International Journal of Computer Applications, 165*(2), 28–35. https://doi.org/10.5120/ijca2017913800

[19] Gheyas, I. A., & Smith, L. S. (2011). A novel neural network ensemble architecture for time series forecasting. *Neurocomputing, 74*(18), 3855–3864. https://doi.org/10.1016/j.neucom.2011.08.005

[20] Tripathi, K. C., & Das, I. M. L. (2009). Teleconnection between Antarctic sea ice concentration and all India rainfall. *Proceedings of International Symposium on Snow and Avalanches*.

[21] Guhathakurta, P., & Rajeevan, M. (2008). Trends in the rainfall pattern over India. *International Journal of Climatology*. https://doi.org/10.1002/joc.1640

[22] Charney, J. G., & Shukla, J. (1981). Predictability of monsoons. In J. Lighthill & R. P. Pearce (Eds.), *Monsoon dynamics* (pp. 99–108). Cambridge University Press.

[23] Tripathi, K. C., Das, M. L., & Sahai, A. (2006). Predictability of sea surface temperature anomalies in the Indian Ocean using artificial neural networks. *Indian Journal of Marine Sciences, 35*, 210–220.

[24] Verma, A., Khadke, L., & Budakoti, S. (2025). Understanding the dynamics of 2024 extreme heat event in India: Spatial variability, hydrometeorological impacts, and model evaluation. *Atmospheric Research, 322*, 108154. https://doi.org/10.1016/j.atmosres.2025.108154

[25] Mohammed, A., & Kora, R. (2023). A comprehensive review on ensemble deep learning: Opportunities and challenges. *Journal of King Saud University - Computer and Information Sciences, 35*(2), 757–774. https://doi.org/10.1016/j.jksuci.2023.01.014

MANAV RACHNA | NAAC A ACCREDITED

Swami Vivekananda Library and Resource Centre
Manav Rachna University

SIMILARITY VERIFICATION CERTIFICATE

Title: Local & Global Prediction Algorithms : The effectiveness in Short / Long range forecasting of extreme Meterological events in Indian subcontinent

Researcher: Rashi Aggarwal Registration Number: MRUCSDN012K210053

Supervisor: Rashi Aggarwal Department: MRUCSDN012K210053

This is to certify that the above mentioned Short Synopsis/Thesis was scanned for similarity index through Drill BIT.....Software on 16.5.25..... with submission ID 3576911... and the similarity index was found to be ...4.....%. The complete report running into .23. Pages, 5173....Word count with File Size 1563.....KB is being forwarded to the Research Supervisor/DRC/DRB for review.

Date: 16.5.25

Swami Vivekananda Library and Resource Centre, Faridabad ★ Manav Rachna University ★
16/5/2025

The complete report of the above Short Synopsis/Thesis has been reviewed and the similarity index including AI-based writing tools are below the accepted norms as per the University Grants Commission (Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions) Regulations 2018. The Short Synopsis/Thesis may be considered for further processing

Rashi
Research Scholar

Manpreet Kaur
Research Supervisor

Dr. Manpreet Kaur
16/5/25

Swami Vivekananda Library and Resource Centre, Faridabad
Manav Rachna University
16/5/2025

Lara Manpreet Kaur
16/5/25

DrillBit

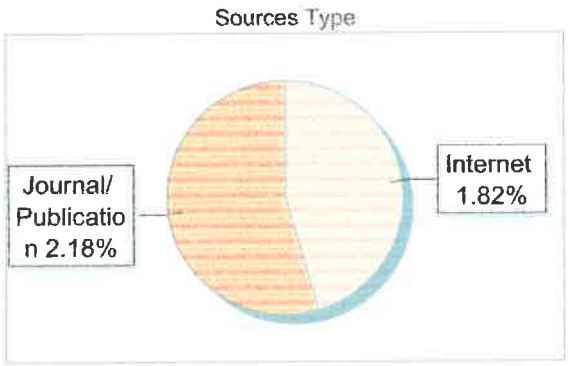
The Report is Generated by DrillBit Plagiarism Detection Software

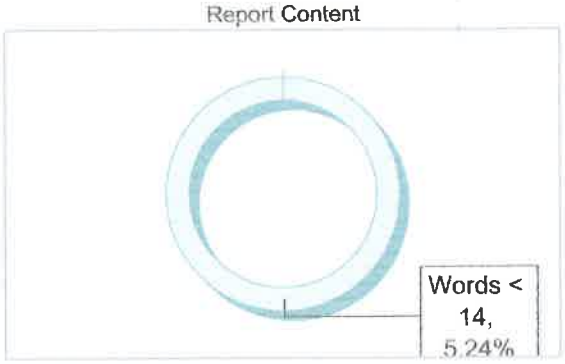
Submission Information

Author Name	Rashi Aggarwal
Title	LOCAL AND GLOBAL PREDICTION ALGORITHMS: THE EFFECTIVENESS IN SHORT/ LONG RANGE FORECASTING OF EXTREME METEOROLOGICAL EVENTS IN THE INDIAN SUBCONTINENT
Paper/Submission ID	3576911
Submitted by	netrapal.library@mru.edu.in
Submission Date	2025-05-05 13:33:15
Total Pages, Total Words	23, 5173
Document type	Synopsis

Result Information

Similarity 4 %





Exclude Information

Quotes	Excluded
References/Bibliography	Excluded
Source: Excluded < 14 Words	Excluded
Excluded Source	0 %
Excluded Phrases	Not Excluded

Database Selection

Language	English
Student Papers	Yes
Journals & publishers	Yes
Internet or Web	Yes
Institution Repository	Yes

Excluding matches
less than 14 words

DrillBit

DrillBit Similarity Report

Manpreet Kaur
16/5/25

4	7	A
SIMILARITY %	MATCHED SOURCES	GRADE

A-Satisfactory (0-10%)
B-Upgrade (11-40%)
C-Poor (41-60%)
D-Unacceptable (61-100%)

LOCATION	MATCHED DOMAIN	%	SOURCE TYPE
1	translate.google.com	1	Internet Data
2	www.arxiv.org	1	Publication
3	[illegible]	1	Publication
4	Liquid chromatographytandem mass spectrometry applications in endocrinology by Mar-2009	<1	Publication
5	drishtiias.com	<1	Internet Data
6	Statistical downscaling of pattern projection using multi-model output variables by Hongwe-2011	<1	Publication
7	worldwidescience.org	<1	Internet Data

Swami Vivekananda Library and Resource Centre, Faridabad
Manav Rachna University
16/5/2025

DrillBit

The Report is Generated by DrillBit AI Content Detection Software

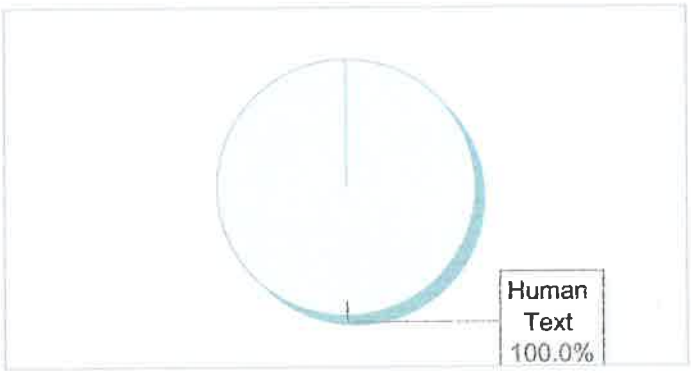
16/5/25

Submission Information

Author Name	Rashi Aggarwal
Title	LOCAL AND GLOBAL PREDICTION ALGORITHMS: THE EFF..
Paper/Submission ID	3576911
Submitted By	netrapal.library@mru.edu.in
Submission Date	2025-05-05 13:33:15
Total Pages	23
Document type	Synopsis

Result Information

AI Text: 0 %



Disclaimer:

* The content detection system employed here is powered by artificial intelligence (AI) technology.
* Its not always accurate and only help to author identify text that might be prepared by a AI tool.
* It is designed to assist in identifying & moderating content that may violate community guidelines/legal regulations, it may not be perfect.